On modal availability, travel strategies and traffic equilibrium on a multimodal network

Fabien Leurent

On modal availability, travel strategies and traffic equilibrium on a multimodal network

Fabien Leurent ([1])
Université Paris Est, Laboratoire Ville Mobilité Transport, Ecole des Ponts ParisTech

Abstract

Transportation modes, including Walking and other private modes as well as transit services, provide travel options to the individual trip-maker along a transportation network. On a single basis, any modal option is featured here in terms of travel time to destination conditionally to immediate availability, expected wait time in the adverse case and the probability of availability. The paper is focused on availability in order to state its roles in travel strategy for route choice and traffic assignment onto a multimodal network. It deals with, successively, mode characterization, local travel strategy, network hyperpaths and traffic equilibrium. Assumedly, a private mode is available on a full, continuous basis, while a transit service is available on a partial, discrete basis due to station dwell time and service frequency. However, a capacitated transit service under saturation amounts to a fully available travel option which includes an initial wait time. A local travel strategy at a choice node is made up of either an ordered sequence of discrete options, or a continuous option, or a combination of both so that the discrete options are used opportunistically if available or the continuous one otherwise. This leads us to revisit the common line problem of transit assignment. A framework and an algorithm are provided to search for optimal travel strategies. The sequential treatment along a multimodal network is based on hyperpaths under availability conditions. Traffic equilibrium is addressed in the static setting; the system state includes the vector of trip flows by destination and network link together with vectors of strategy proportions by node and destination.

Keywords

Multimodal assignment. Traffic equilibrium. Hyperpaths. Capacitated assignment.

1. Introduction

Background. In the theory of traffic assignment to a transportation network, the concept of travel strategy pertains to the route choice behavior of a network user under dynamic information about traffic conditions. Put simply, a travel strategy is a set of local travel options (i.e. modes or paths) which the user is willing to take in order to get closer to his destination. The concept has been designed initially for urban bus networks, along which several lines may be attractive jointly from a given station to get to a destination. In order to generalize the common line problem of Chriqui and Robillard [ChR75] to the network setting, Spiess and Florian [SpF89] assumed that when the transit user arrives at a stop then he waits for the first attractive vehicle to come. On assuming independent, Markovian services, each attractive line is assigned a share of flow that is equal to the ratio of its frequency to the combined frequency of all attractive lines.

[1] Corresponding author: fabien.leurent@enpc.fr. Ecole des Ponts ParisTech, Laboratoire Ville Mobilité Transport, 6-8 avenue Blaise Pascal, Champs sur Marne 77455 Marne la Vallée Cedex

The concept of strategy has been elaborated in several directions to model more specific situations and/or dynamic setting. For instance, [HMN04] added a notion of ordering to assign a trip flow in compliance with capacity constraints, whereas [CoC01] addressed congestible strategies by relating the local frequency to the residual capacity.

Objective. However the presence or absence of a given service when the user arrives at a choice node, constitutes an eventuality which has not been modeled so far, to the best of our knowledge (except for our own attempt in a French textbook, [Leu06]). From this eventuality stems the issue of the immediate availability of a given transportation mode: which option is preferable out of those immediately available at a given instant, or is it better to wait for another option to become available? The paper provides a framework to represent availability in the setting of local route choice (i.e. local travel strategy) and also in the setting of network hyperpaths and traffic equilibrium. For simplicity, system state is taken stationary. Multimodal travel services are considered, including Walking and other private modes (e.g. by car) together with public services notably transit lines and maybe also Vehicle-Sharing systems.

Contribution. Immediate availability of a given travel option (e.g. mode or path) at a given place is modeled by its probability at the instant of user arrival, the modal Probability of Immediate Availability (PIA). The average travel cost of that option involves the PIA and the eventual cost depending on whether or not the option is immediately available. Independence between availability events is assumed between different modes and between different forms of a given mode. At a choice node, a travel strategy is a subset of the travel options supplied there, involving either partially available services only or a single deterministic option or a hybrid strategy. There is an order of consideration to deal with multiple options within a strategy. A problem of optimal travel strategy is stated and an efficient solution algorithm is supplied, which bears strong similarities to, yet significant differences from, the common line problem.

In the network setting, local travel strategies are chained sequentially into hyperpaths. The problem of shortest hyperpath under availability conditions is stated; it is shown to have a solution of which the cost is unique. Furthermore, shortest hyperpaths targeted to a given destination have a recursive structure, so that they satisfy a set of generalized Bellman equations. Ford-Bellman and Dijkstra-like algorithms are provided to search for shortest hyperpaths. The network loading of origin-destination flows along a hyperpath under availability proceeds in the conventional way. Lastly, the issue of traffic equilibrium is addressed by modeling the sensitivity to trip flows of the local traffic conditions – namely the PIA, the minimum link time and service frequency. Conditions of user optimality are stated to define traffic equilibrium. It is characterized by a Variational Inequality with respect to a vector of strategy assignment by node and destination. Under mild assumptions the cost mapping is continuous and the feasible set is nonempty: therefore an equilibrium must exist. A Method of Successive Averages is recommended as a heuristic equilibration algorithm, together with a rigorous convergence criterion.

Paper organization. The rest of the paper is in five parts. Section 2 provides the basic model of a modal option, with discussion of the potential instances. Section 3 deals with local travel strategies: after defining a strategy and its cost to a user, the problem of optimal strategy is stated and solved by an efficient algorithm. Then, Section 4 addresses the network setting by providing definition of, and algorithmic treatment for, hyperpaths under availability conditions. Next, Section 5 brings about the traffic equilibrium framework. Lastly, Section 6 deals with a numerical instance.

Table of notation

N set of nodes indexed by n or m

A set of links $a \approx (n,m)$ with tail and head nodes $n,m \in N$

A_n^+ (resp. A_n^-) set of links going out (resp. coming in) node n

ρ_a probability of immediate availability of link a

t_a^o minimum travel time

φ_a service frequency (if a transit link)

s travel strategy at node n, $S(n)$ set of such strategies, split into subset S_D of deterministic options, S_O of ordered sequences and S_H of hybrid sequences

C_s (resp. T_s) average cost (resp. recourse cost) of strategy s

π_a^s proportion of link a within strategy s

2. Mode characterization

2.1 Core assumptions

Let us consider an individual trip-maker at a given place, say node n, in a transportation network, where a set A_n^+ of modal options a are supplied to him. By assumption, option a is available immediately with probability ρ_a: yielding then a time to destination of t_a^o. Otherwise, the trip-maker has to wait a time length of w_a for the option to become available. Overall, the time to destination is a random variable, t_a, with average value

$$\bar{t}_a = \rho_a t_a^o + (1-\rho_a)(t_a^o + \mathrm{E}[w_a]) = t_a^o + (1-\rho_a)\overline{w}_a \,. \tag{1}$$

More generally, generalized travel time including the money expense as well as the time expense can be modeled. To each kind of time, at waiting versus at running, is associated a specific value to take into account its comfort and its reliability.

Private modes such as Walking or the Private Bike or the Private Car can be taken as continuously available, with $\rho_a = 1$ and $w_a = 0$. A public mode such as the Cab or a Sharing-Vehicle System has its availability restricted to a probability less than one: typically equal to the load ratio in an elementary Markovian queuing system (e.g. [Kle75]). A transit service with frequency of φ_a runs during period H and useful dwell time of δ_a by run is immediately available with probability $\rho_a = \min\{1, \delta_a\varphi_a / H\}$.

In this setting, the average time of a transit option, $t_a^o + (1-\rho_a)\overline{w}_a$, is inferior to that commonly taken in static transit assignment where availability is neglected, which would correspond to $\rho_a = 0$. Indeed, high frequency services with φ_a about 30 or even 40 trains per hour and usable dwell time of say 40 seconds have an availability probability of 1/3 or 4/9. Then factor $1-\rho_a$ is significantly less than 1; however it is applied to an average wait time that is quite short, so that the conventional approximation is not so bad concerning average time.

2.2 Survival time

If the travel option is not available at the instant of user arrival, then the time to wait until it becomes available is a residual time. For a transit option, the user wait time on the station platform amounts to the residual span (residual lifetime) of the on-going headway interval – defining the headway as the interval between two instants of successive vehicle departure. From survival theory (e.g. [Kle75]), its Probability Density Function $\dot{\mathrm{W}}_a(x)$ stems from the Cumulative Distribution Function $\mathrm{H}_a(x)$ of the headway η_a in the following way:

$$\dot{\mathrm{W}}_a(x) = [1 - \mathrm{H}_a(x)]/\mathrm{E}[\eta_a]. \tag{2}$$

Then it holds that $\mathrm{E}[w_a^k] = \mathrm{E}[\eta_a^{k+1}]/(k+1)\,\mathrm{E}[\eta_a]$, yielding average user wait time of

$$\overline{w}_a = \mathrm{E}[\eta_a^2]/\,2\,\mathrm{E}[\eta_a]. \tag{3}$$

In the Markovian case, headway η_a is distributed exponential and so is the associated wait time w_a, with the same parameter φ_a / H. But if the transit service is perfectly regular, then $\eta_a = H/\varphi_a$ is deterministic, yielding a uniform distribution of wait time on $[0, H/\varphi_a]$ with average $\overline{w}_a = H/2\varphi_a$ and CDF $\mathrm{W}_a(x) = \min\{1, x\varphi_a / H\}$ for $x \geq 0$. Irregular services yield still more variable wait times.

2.3 Capacity constraints

Both the vehicle capacity of road or rail infrastructure and the passenger capacity of public modes are limited. Denoting by κ_a the capacity rate of link a during the period of interest and by v_a the trip rate, it must hold that $x_a \leq \kappa_a$. When $x_a < \kappa_a$, there is no saturation and it is not necessary to consider the limitation. However, in the saturated case it is crucial to understand its features and consequences: a prominent feature is the persistence of a user stock, denoted as N_a, which does not vanish during the period. Here it is taken as a stable variable, despite it will fluctuate with the customer arrivals and the delivery of capacity. For a private mode under priority queuing, to get from the back to the front of the stock requires a wait time of $w_a = N_a / \kappa_a$.

For a transit service, let us assume that each vehicle run supplies a capacity of k_a passengers at the station: then $\kappa_a = k_a\,\varphi_a / H$. Under priority queuing, a passenger is likely to wait for the n-th run after his instant of arrival to succeed to board, with $j = 1 + \mathrm{Int}(N_a / k_a)$: his wait time is the sum of $j-1$ service headways plus the initial survival time: so $\overline{w}_a^{\mathrm{Q}} = (j-1)\overline{\eta}_a + \overline{w}_a$.

Alternatively, waiting passengers may be mingling rather than queuing: then, according to [KBS03], there is a probability of immediate boarding, p_a, for each user in each vehicle run, so that the rank of the available vehicle counted from the instant of user arrival is a binomial random variable with parameter p_a, yielding an average wait time of $\overline{w}_a^{\mathrm{M}} = \overline{w}_a + \overline{\eta}_a / p_a$. As shown by [LeC012], the probability of immediate boarding must stem from the passenger stock as follows: $p_a = \min\{1, k_a / N_a\}$.

Whatever the waiting discipline, the event of immediate boarding is definitely distinct from immediate availability. To avail oneself of the transit service under saturation, the user has to join the stock, thus making a decision and choosing his route. Therefore, as a travel option, a public service under saturation must be considered as immediately available, with $\rho_a = 1$ and travel time $t_a^o + \overline{w}_a^{\mathrm{Q/M}}$.

2.4 Dynamic vs. prior information

Let us make a distinction between prior information that is known to a recurrent trip-maker prior to arriving at a given place – e.g. estimated time to destination or nominal frequency – and dynamic information that pertains to variable conditions, which it depicts more accurately. By nature, partial availability is a variable condition associated to random events of availability or unavailability: dynamic information about the current status does hold at the instant of user arrival. In transit assignment, it has long been assumed that no vehicle is dwelling when the trip-maker arrives at the station and that his dynamic information consists in witnessing to the next arrival of a service vehicle and identifying the associated line. Here it is assumed that the user is aware of availability. As concerns prior information, we assume that the user has a notion of the travel options, each one with its probability of immediate availability (PIA) and the expected time to destination depending on the availability status. The size of the passenger stock at a transit station also constitutes an element of information that is easy to estimate crudely on arriving at the station platform.

3. Local route choice and travel strategies

Let us now consider a situation of route choice as follows: at a given place, the user is supplied with several travel modes or routes within a given mode. He has to select one option to get closer to his destination. He will prefer an option of minimum cost to him, according to the axiom of actor rationality in microeconomic theory. However the outcome will depend on the current system state at his instant of arrival, which includes the status of availability of the choice options, and also on the system evolution if no attractive option is immediately available. Thus the routing behavior is featured as an adaptive choice process.

3.1 Travel strategies

Here a travel strategy is defined as a usage pattern for a subset of travel options belonging to one out of the following three kinds. Firstly, a deterministic option that is fully available: this is the deterministic kind. Secondly, an ordered sequence of partly available options: the user will select the first option that is immediately available in the sequence, or otherwise the first one to become available: this is the discrete sequence kind. Thirdly, a combination of a discrete sequence and a deterministic option called the recourse option: assumedly, the user will select the first option that is immediately available within the discrete sequence, or otherwise the recourse option. This is the hybrid kind of strategy.

The definition ensures that any option gives rise to a singleton strategy and can be involved into bundle strategies. It emphasizes the issue of availability. The order in a discrete sequence is crucial to avoid any ambiguity, as it may happen that several discrete options be available jointly. Indeed, it is essential to describe the system state and its evolution in a consistent way by taking into account the various eventualities of the joint status of availability. Denoting by s an ordered discrete sequence and indexing by i the options in $s = [a_i : i \in \{1,..\bar{s}\}]$, then the first option is immediately available with probability ρ_{a1}, which makes its basic probability of selection within s, denoted r^s_{a1}. Then the immediate availability of the second option is of interest to the user in a proportion $\bar{\rho}_{a1} \equiv (1-\rho_{a1})$ of the cases, on assuming that the discrete options are mutually independent: its basic probability of selection is $r^s_{a2} = \bar{\rho}_{a1}.\rho_{a2}$. By induction, denoting $R^s_i \equiv \sum_{j=1}^{i} r^s_{a(j)}$, the next option has basic selection probability of

$$r^s_{a(i+1)} = (1 - R^s_i).\rho_{a(i+1)}\,. \tag{4}$$

The residual probability, $\overline{R}_s \equiv 1 - R^s_{\overline{s}}$, is the probability of joint immediate unavailability. For a hybrid strategy it is the probability of the recourse, deterministic option. For a discrete sequence, the recourse option is to wait for the first option to become available. Assuming independent Markovian services, the wait time for option a is an exponential random variable w_a with parameter φ_a / H. Then, the wait time for the next attractive run, $w_s \equiv \min_{a\in s} w_a$, is an exponential random variable with parameter φ_s / H, in which $\varphi_s = \sum_{a\in s}\varphi_a$ denotes the combined frequency. Furthermore, the probability of option a being the first to become available in the recourse case amounts to φ_a / φ_s.

Thus, the total probability of option a within discrete sequence s (its market share) amounts to

$$\pi^s_a = r^s_a + \overline{R}_s\, \varphi_a / \varphi_s\,. \tag{5}$$

This property of the travel strategy is not likely to be known to the user, who merely applies his predefined order to adapt oneself to the current system state.

By the way, two assumptions of independence between partly available options and of Markovian services have been made. They are required to obtain tractable formulae and provide a stochastic background to the model, which may lack realism but suffices to depict stochastic variability and to trace its consequences.

3.2 Strategy cost

The cost of a travel strategy is defined as the average expected cost to its user over all eventualities. This involves the probabilities of the underlying random events associated to the choice options that belong to the strategy. By kind of strategy the cost is as follows:

- a deterministic option $s = \{a\}$ yields its average cost, $C_s = t^o_a$.
- A discrete sequence yields either minimum option cost $t^o_{a(i)}$ with basic option probability $r^s_{a(i)}$ or a combined cost including waiting, $(\tilde{H} + \sum_{a\in s} t^o_a \varphi_a)/\varphi_s$, with recourse probability $\overline{R}_s$. So the expected cost is $C_s = \sum_{a\in s} t^o_a\, \pi^s_a + \overline{R}_s \tilde{H} / \varphi_s$.
- In a hybrid strategy, the recourse option b has deterministic cost t_b, so the expected cost is $C_s = \sum_{a\in s} t^o_a\, r^s_a + \overline{R}_s\, t_b$.

Notation H has been adapted to $\tilde{H} = \gamma H$ in order to penalize wait time differently from run time, or to accommodate non-Markovian services by using the approximation $\overline{w}_a = \gamma H / \varphi_a$ (e.g. $\gamma = \frac{1}{2}$ for regular services).

3.3 The optimal strategy problem

The problem of optimal strategy is to find a strategy of minimum cost within the set of all admissible strategies, $S = S_D \cup S_S \cup S_H$. Assuming a finite and nonempty set of options, then the admissible set is finite and nonempty so there must be a solution. If there is a deterministic option of which the cost is less than those of the other deterministic options and than the minimum costs of the partly available options, then the associated singleton strategy is optimal. Conversely, if there is a partly available option, say a, with minimum cost less than that of any deterministic option b, then the hybrid strategy $\{a,b\}$ is better than strategy $\{b\}$. Furthermore, if $t_a^o + (1-\rho_a)\tilde{H}/\varphi_a$ is less than t_b then $\{a\}$ as a discrete sequence is less costly than $\{b\}$.

Let us define the recourse cost of a strategy s as either $T_s \equiv t_a^o$ if $s = \{a\}$ is deterministic, or the combined cost at unavailability if s is an ordered sequence, or as the cost of the deterministic option in a hybrid strategy. Then, let us define the relationship of attractiveness of an option a to a strategy s by the characteristic property as follows: $t_a^o \leq T_s$.

Proposition 1*. *Structural properties of attractiveness and optimality:

(i) *A partly available (PA) option a is attractive to an ordered sequence s iff any ordered sequence say s' made of the insertion of a in any position in s has $T_{s'} \leq T_s$.*

(ii) *Any ordered sequence s including two options a and b at rank i and $i+1$ respectively such that $t_a^o > t_b^o$ can be rearranged into s' by swapping those options, yielding reduced strategy cost.*

(iii) *The cost of any ordered sequence is reduced by rearranging its options in order of increasing minimum time.*

(iv) *A deterministic option a that is attractive to an ordered sequence s gives rise to a hybrid option $s' \equiv s \cup a$ that is cheaper than s.*

(v) *Any hybrid strategy (resp. ordered sequence) can be reduced by dropping its PA options of minimum time greater than its recourse cost, yielding reduced strategy cost.*

Proof. (i) Whatever the position of insertion for option a into list s, the extended strategy has recourse cost of $T_{s'} = (\varphi_a t_a^o + \varphi_s T_s)/\varphi_{s'}$. Then, the condition that $T_{s'} \leq T_s$ is equivalent to $\varphi_a t_a^o + \varphi_s T_s \leq \varphi_{s'} T_s$ hence to $\varphi_a t_a^o \leq \varphi_a T_s$: thus, if a is attractive to s then $T_{s'} \leq T_s$ and conversely.

(ii) Such swapping has no effect on the combined cost under unavailability. Nor does it change the joint probability of choice under availability, $r_i^s + r_{i+1}^s$, nor the residual probability, $1 - R_{i+1}^s$, since respectively

$$r_i^s + r_{i+1}^s = (1 - R_{i-1}^s)\rho_a + (1 - R_{i-1}^s)(1-\rho_a)\rho_b = (1 - R_{i-1}^s)(\rho_a + \rho_b - \rho_a\rho_b),$$

$$1 - R_{i+1}^s = (1 - R_{i-1}^s) - r_i^s - r_{i+1}^s.$$

Thus the only change in strategy cost between s and s' pertains to the replacement of $r_i^s t_a^o + r_{i+1}^s t_b^o$ by $r_i^{s'} t_b^o + r_{i+1}^{s'} t_a^o$ with increased share of total weight $r_i^s + r_{i+1}^s$ assigned to the cheaper option, thus yielding a cheaper union.

(iii) stems from (ii) by obvious induction.

(iv) By reducing the recourse cost, Coeteris paribus.

(v) In this case the more expensive PA options, used when cheaper options are not available, are replaced by the cheaper recourse option, so the strategy cost is diminished.

Proposition 2. *One of the following three cases must hold:*

(i) *a deterministic option has overall minimum cost, yielding optimal deterministic strategy.*

(ii) *An ordered sequence strategy has recourse cost minimum among S_O and less than any deterministic cost, yielding optimal discrete strategy after ranking its options in order of increasing minimum time.*

(iii) *An ordered sequence strategy has recourse cost minimum among S_O and all of its minimum costs less than any deterministic cost, but its self-recourse cost higher than at least one deterministic cost: then any hybrid of that sequence ordered by increasing minimum travel time, with a deterministic option of minimum cost, is globally optimal.*

Proof. (i) Since then no PA option can improve on the minimum deterministic cost, hence no ordered sequence nor hybrid strategy. Cases (ii) and (iii) are complementary to (i) and mutually exclusive (save for cost equality between the self-recourse option of an optimal ordered sequence and the minimum cost among deterministic options). In (ii) and (iii) the requirement to rank the options stems from Proposition 1(ii).

3.4 Optimal strategy algorithm

Let us divide the set of options in two subsets of, respectively, partly available (PA) options and fully available ones (FA). Then, let us rank the PA options in order of increasing bare time and the FA set in order of increasing cost. Denote by c_{PA} (resp. c_{FA}) the minimum cost in each subset.

If $c_{\text{FA}} \le c_{\text{PA}}$ then any FA option of minimum cost yields an optimal strategy. Otherwise, determine the optimal ordered sequence strategy along the following steps, based on the list $[a_i : i \in \{1,..I\}]$ of the I PA options ordered by increasing minimum cost:

a) Let $i := 1$, $r_1 := \rho_{a(1)}$, $R := r_1$, $B := r_1 t^o_{a(1)}$, $F := \varphi_{a(1)}$, $W := \varphi_{a(1)} t^o_{a(1)}$.

b) Let $C := B + (1-R).(\tilde{H} + W)/F$. If $i = I$ or $C < t^o_{a(i+1)}$ then Terminate with ordered sequence $s_i := [a_j : j \in \{1,..i\}]$, else let $i := i+1$ and continue.

c) Let first $r_i := (1-R).\rho_{a(i)}$, then $R := R + r_i$, $B := B + r_i t^o_{a(i)}$, $F := F + \varphi_{a(i)}$, and $W := W + \varphi_{a(i)} t^o_{a(i)}$. Go to Step b.

The auxiliary variables are purported to accumulate, respectively, probability of events (R), frequency of agreed PA options (F), minimum time weighted by availability probability (B), minimum time weighted by frequency (W). The process takes end when no other PA option can be included to reduce the combined cost C.

If the final recourse cost, $(\tilde{H} + W)/F$, is less than c_{FA}, then no FA option can contribute to improve on the optimal ordered sequence. Conversely, if $c_{\text{FA}} < (\tilde{H} + W)/F$ where $c_{\text{FA}} = t^o_b$ of FA option b, then the hybrid strategy $s_i \cup b$ has a cost inferior to that of s_i. This entails that only a hybrid strategy can be optimal: it must include b and also any PA option $a(j)$ in s_i such that $t^o_{a(j)} < t^o_b$. An option $a(j)$ such that $t^o_{a(j)} = t^o_b$ can be included in the optimal strategy, with no consequence on the cost but some effect on the market shares. Any option

$a(j)$ in s_i with $t^o_{a(j)} > t^o_b$ would deteriorate the combined cost and cannot take part to the optimal hybrid.

To ease the hybridization, in Step b of the algorithm it suffices to replace the formula for C by the following

$$C := B + (1-R).\min\{t^o_b, (\tilde{H}+W)/F\}.$$

At termination, if $t^o_b \geq (\tilde{H}+W)/F$ then the optimal strategy is an ordered sequence, otherwise it is hybridized with FA option b.

The computational complexity of the Optimal Strategy Algorithm (OSA) amounts to $\mathrm{O}(\overline{a}\ln(\overline{a}))$ with $\overline{a}$ the number of options, since each option is dealt with in $\mathrm{O}(1)$ elementary operations, except for the initial ordering which can be accomplished in $\mathrm{O}(\ln(\overline{a}))$ by option.

Proposition 3*. The revised algorithm yields an optimal strategy.*

Proof. Because either there are only deterministic options or also PA options but with larger minimum cost, yielding case (i) in Prop. 2, or there are PA options of minimum time less than any deterministic cost. In the latter case, the algorithm yields an ordered sequence s that is optimal among S_O because each of its constituents takes an active part in minimizing the strategy cost, whereas no outside PA option is attractive with respect to s. Furthermore, it is ordered correctly. Lastly, if a deterministic option b with $t^o_b = c_{\mathrm{FA}}$ is attractive to s, then the hybrid $s \cup b$ is globally optimal.

3.5 The common line problem revisited

The common line problem of [ChR75] is a problem of optimal strategy restricted to transit lines and neglecting the issue of availability. Then, it is not required to order the options in a strategy. The associated algorithm amounts to the OSA with $\rho_a = 0$.

The consideration of availability induces two noticeable changes. First, the cost of a given strategy is reduced by including the event of availability, thereby diminishing the probability to wait and the expected wait cost. So a line which would be attractive in the model without availability may not be so in the availability model. In the binary case, this amounts to the twofold condition that $t^o_a + (1-\rho_a)\tilde{H}/\varphi_a < t_b < t^o_a + \tilde{H}/\varphi_a$.

Second, in the network setting [SpF89] it has become customary to include deterministic options in the common line problem, by assimilating them to transit services of same minimum time and with infinite frequency. Under the convention that $\infty/\infty = 1$, the cost formula applied to a combination of lines yields only the deterministic time t_b, whatever the value of the minimum times of the transit lines, and the formula is applied with $t_b < C$, the expected wcost of the strategy under optimization. Therefore, transit lines with $t^o_a < t_b < C$ are neglected, whereas they get a positive market share in the availability model.

4. Network hyperpaths under availability conditions

In this section, the local travel strategies are taken into the perspective of a multimodal transportation network so as to define network paths for any trip between any origin-destination pair of network nodes. After providing network definition and notation (§ 4.1), the notion of hyperpath is adapted to the availability framework (§ 4.2) and a solution is shown to exist (§ 4.3). Then, two shortest path algorithms respectively à la Ford-Bellman and à la Dijkstra-Spiess-Florian are provided (§ 4.4).

4.1 Network setting

Let us define the transportation network as an oriented graph $[N, A]$ with N the set of nodes (vertices) and A the set of links (oriented arcs) $a \approx (n, m)$ between nodes $n, m \in N$. A node represents a given place, eventually with one or several links going out of it thus providing opportunity for route choice. A link represents a transition from its tail node to its head node: it must belong to a given travel mode such as Walking or Transit or Vehicle sharing. A link with public mode has one function only out of the following three ones: access (or boarding), egress (or alighting) or transit. If it belongs to a transit mode then it must be associated with a given line service.

The network is consistent with the modal specification if any public link with transit or exit function can be accessed only from a public link of the same mode (resp. line service) with function of access or transit and if any public link with access or transit function yields access only to a public link of the same mode (resp. line service) with function of transit or egress.

Set A_n^+ (resp. A_n^-) encompasses the links that end in (resp. go out of) node n. Node set Z brings together the destination places.

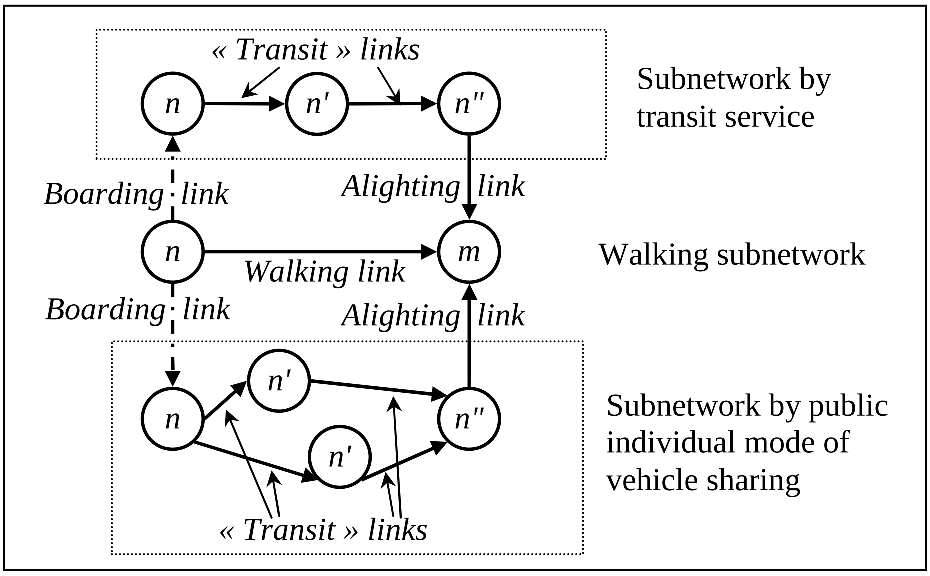

Fig. 1. Topology of nodes, links and modes.

4.2 From local strategy to network hyperpaths

A hyperpath to a specified destination node z is defined as an acyclic sub-graph providing paths to its destination from any of its incident nodes save the destination. This definition is taken from Nguyen and Pallotino (1988) who applied it to transit networks with a set of transit lines and a Walk mode. Conventionally, a given hyperpath h is specified by associating to each node $n \in N$ a subset $A^+_{n/h}$ of the links that go out of n and belong to h. It is the basic topological structure to accommodate origin-destination flows consistently with dynamic state at any choice place, by taking subset $A^+_{n/h}$ as a local travel strategy. Then, the acyclic structure of the hyperpath determines the serial combination of local strategies along any path sequence and the associated sequence of local routing probabilities yields the probability of that path along the hyperpath.

Under the availability framework, the notion of hyperpath must be adapted in order to accommodate the specific nature of the admissible travel strategies. Denoting by FA_n^+ and PA_n^+ the sets of links emanating from node n that have full or partial availability, respectively, an admissible strategy s_n must include at least one link in A_n^+ and at most one link in FA_n^+. Furthermore, if $s_n \cap PA_n^+$ contains more than one link, then the order among them should be specified, as well as that between $s_n \cap PA_n^+$ and $s_n \cap PA_n^+$ if both are nonempty. Let us address the latter issue by assuming that it corresponds to a hybrid strategy in which the FA option is the recourse one after all the PA options have been included. To deal with the former issue of sequential ordering, let us further specify an admissible strategy by assuming that the part in PA_n^+ is an ordered list.

Overall, this definition of local strategies induces a larger set of network hyperpaths; for any finite set A_n^+ it is still finite since there is a finite number of orderings within any subset of PA_n^+, and the number of such subsets is finite.

The definitions provide unambiguous local routing probabilities along a given hyperpath h: $\forall n \in h \setminus z$, any link in A_n^+ has a market share π_a^h equal to 0 if $a \notin h$ or else to π_a^s from the local strategy. Then, any path r along the hyperpath has a market share (routing proportion) between his origin and the destination of $\pi_r^h = \prod_{a \in r} \pi_a^h$.

Therefore the basic properties of flows along a hyperpath hold, notably so the recursive flow loading of origin-destination flows by a forward algorithm dealing with the hyperpath nodes in topological order towards the destination.

Travel cost along hyperpath h between any node n taken as origin and destination z is evaluated recursively, starting from z and dealing with each node taken in reverse topological order. When node n is addressed, its downstream nodes m along h have already been dealt with, yielding costs g_{mz}^h: thus each link $a \approx (n,m) \in h$ has a minimum travel cost up to destination of $t_a^{oh} \equiv t_a^o + g_{mz}^h$ and an average cost of $t_a^{oh} + \overline{w}_a$, eventually associated to a probability of immediate probability, ρ_a, and a service frequency φ_a if it is a transit link.

This enables one to obtain unambiguously the hyperpath cost from n along h, namely

$$g_{nz}^h = \mathrm{C}_{nz}(s, [t_a^{oh}, w_a, \varphi_a : a \approx (n,m) \in A_n^+]). \tag{6}$$

Notation $\mathrm{C}_{nz}(s,.)$ stands for the function of local strategy cost as stated in Section 3.2.

4.3 The shortest hyperpath problem subject to availability conditions

The problem of shortest hyperpath between nodes n and z is to find a hyperpath destined to z with minimum cost from n.

***Lemma 1*. *Existence of a shortest hyperpath*.** *If the set of hyperpaths between nodes n and z is nonempty, i.e. if there is a path from n to z on the network, then there is a shortest hyperpath between the two nodes.*

Proof. Under the feasibility condition, the problem has a solution because the admissible set is nonempty and finite: a finite network gives rise to a finite number of acyclic subgraphs and at each node in such a subgraph there is a finite number of orderings for any subset of its outgoing links.

The availability condition may disrupt a property that holds for basic hyperpaths: namely, the reduction of node cost to destination along a hyperpath. For instance, let us take a three link network as depicted in figure 2. Transit link $a \approx (n,z)$ has minimum cost $t_a^o = 9$', average wait $\overline{w}_a = 6$' for Markovian service of frequency $\varphi_a = 10$ per hour and a probability of immediate availability $\rho_a = .2$. Walk link $d \approx (m,z)$ has cost $t_d = 11.8$' and walk link $b \approx (n,m)$ has cost $t_b = .2$'. The optimal cost from m to z is $c_{mz}^* = 11.8$'. But at node n the optimal strategy is a hybrid one with transit link a taken opportunistically and walk link b taken as recourse, yielding average cost of $c_{nz}^* = .2 \times 9 + .8 \times 12$ = 11.4', which is less than $t_a^o + \overline{w}_a = 15$' as well as $t_b + t_d =$ 12', and also less than c_{mz}^* despite that, in topological order, m is closer to z than n.

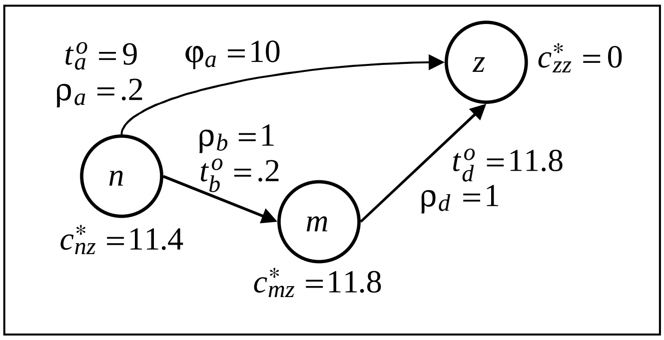

Fig. 2. Small instance.

However, as for basic hyperpaths, a property of recursive structure holds for shortest hyperpaths under availability conditions.

In any hyperpath h destined to z, define the node height u_n^h as the maximum number of links in any path from n to z along h.

Assume now that there exists an oriented path from any node $n \in N \setminus z$ to z.

Proposition 4. *There exists a hyperpath that connects every node $n \in N \setminus z$ to z and which is shortest and lowest among the shortest ones at each of its nodes.*

A straightforward corollary is that a shortest lowest hyperpath from n to z has a recursive internal structure, i.e. $\forall a \approx (n,m)$ its restriction from m is shortest and lowest.

Proof. By induction on the basis of the following inductive property:

$\mathrm{P}[i]$: at order i, there is a hyperpath $\mathrm{H}^{[i]}$ destined to z such that (i) it contains all nodes $n \in N \setminus z$ from which z can be reached by a path of length at most i and (ii) to any node n incident to $\mathrm{H}^{[i]}$ at height at most i, the sub-hyperpath $\mathrm{H}_n^{[i]}$ downstream to n is shortest among the hyperpaths from n to z of length at most i.

At order $i = 1$, $\mathrm{H}^{[1]}$ is obtained by associating to each node n such that $\exists a \approx (n,z) \in A_z^-$ the shortest optimal strategy. The union of the strategies link support sets over such nodes is included in A_z^- so it is a hyperpath and it includes the shortest hyperpaths of length 1.

Let us now assume that the property holds up to order i with associated hyperpath $\mathrm{H}^{[i]}$. At the next order $i+1$, consider any node n connected to z by a path of length $\leq i+1$, say r: denoting by $a \approx (n,m)$ the first link in r, the sub-path $r \setminus a$ has length $\leq i$ so the inductive property at order i applied to m implies that it must be incident to $\mathrm{H}^{[i]}$. Let us build $\mathrm{H}^{[i+1]}$

in the following way: at each node n such that $\exists a \approx (n,m)$ with m incident to $\mathrm{H}^{[i]}$, solve the optimal strategy problem at n restricted to links with head node in $\mathrm{H}^{[i]}$. If n has a shorter hyperpath of length $\leq i$ then it will not improve on the cost in $\mathrm{H}^{[i]}$, due to the inductive property at order i. Otherwise, let us denote by $\mathrm{H}_n^{[i+1]}$ the union of this restricted optimal strategy completed by the downstream hyperpaths $\mathrm{H}_m^{[i]}$. First, $\mathrm{H}_n^{[i+1]}$ must be a hyperpath as the nodes downstream of n all belong to an $\mathrm{H}_m^{[i]}$, which is a hyperpath, and the union of the support sets $\mathrm{H}_m^{[i]}$ contains no loop since the common nodes are involved at same height – their height in $\mathrm{H}^{[i]}$. If n is incident to $\mathrm{H}^{[i]}$ but with a cheaper hyperpath of length $i+1$, its height will change from the value in $\mathrm{H}^{[i]}$ to $i+1$ in $\mathrm{H}^{[i+1]}$. This yields point (i) at order $i+1$. As for point (ii), $\mathrm{H}_n^{[i+1]}$ must be shortest among hyperpaths of height $\leq i+1$ between n and z: if there was a cheaper hyperpath say h of height $\leq i+1$ between n and z, then to have cost strictly less than that of $\mathrm{H}_n^{[i+1]}$ it should include at least one successor node m with cost by h strictly less than by $\mathrm{H}_n^{[i+1]}$ - due to the cost monotonicity in the Optimal Strategy Algorithm. Then, the part h_m downstream of m has length $\leq i$: by the inductive property at order i it cannot have cost lower than that of $\mathrm{H}_m^{[i]}$, thus yielding a contradiction. Lastly, the union of support sets $\mathrm{H}_n^{[i+1]}$ is organized into a global hyperpath since either n has height $\leq i$ so the hyperpath property of $\mathrm{H}^{[i]}$ applies, or it has height $i+1$ as any node that has been newly added or re-positioned at that step. To sum up, the inductive property also holds at $i+1$ if it holds at i, yielding its validity at any order.

The optimality principle can be stated formally as the following system of local conditions on the optimal hyperpath costs $c^*_{m,z}$ from m to z:

$$c^*_{z,z} = 0,$$
$$c^*_{n,z} = \min_{s \in S(n)} \mathrm{C}_{nz}(s, [t^o_a + c^*_{m,z}, w_a, \varphi_a : a \approx (n,m) \in A^+_n]) \tag{7}$$

Wherein C_{nz} has been defined in section 3.2.

4.4 Algorithms for shortest hyperpaths

The two basic algorithms for shortest paths on a network have been extended for shortest hyperpaths of the basic kind: the Ford-Bellman type by Nguyen and Pallotino [NgP88] and the Dijkstra type by Spiess and Florian [SpF89]. Let us adapt them to shortest hyperpaths under availability conditions.

In the Ford-Bellman type algorithm, at each iteration k all the nodes are fetched to re-optimize $c^{(k)}_{n,z}$ on the basis of $c^{(k-1)}_{m,z}$, starting from initial costs $c^{(0)}_{m,z} := \infty$ save for $c^{(0)}_{z,z} := 0$. This ensures that shortest hyperpaths of height k must be discovered at step k. Thus, the Ford-Bellman algorithm under availability is fundamentally equivalent to the constructive proof of Proposition 4. Apart from the specific problem of local optimal travel strategy, another difference from the shortest hyperpath algorithm à la Ford-Bellman in [NgP88] is the requirement to avoid loops by an explicit treatment – which is not necessary in the basic case owing to the monotonicity of cost along a hyperpath of the basic kind. To do that, let us associated to each node n a vector of node indicators to mark its downstream nodes, denoted by $\delta^n = [\delta^n_m : m \in N]$ with $\delta^n_m = 1$ if m is downstream of n or $\delta^n_m = 0$ otherwise. The Ford-Bellman algorithm proceeds as follows, using node costs c_m and local strategies s_m :

Initialization. $\forall m \in N \setminus \{z\}$, let $c_m := \infty$, $s_m := \varnothing$ and $\delta^m_n := 0$ $\forall n \in N \setminus \{m, z\}$ and $\delta^m_z := 1$. Let also $c_z := 0$ and $k := 0$.

Iteration. Let $k = k+1$. For any node m such that $\exists a \approx (m,n)$ such that $c_n < \infty$, do the following. Let s_m be the optimal strategy at m based on the provisional downstream costs $t^o_a + c_n$ and the other modal conditions of link a, under the additional restriction that $\delta^m_n = 0$. This yields strategy cost c'_m. If $c'_m < c_m$ then replace c_m by c'_m and the former local strategy by s_m and δ^m by the union of all δ^n such that $\exists a \approx (m,n) \in s_m$.

Termination test. If $k = \overline{N}$ or no node cost has changed during iteration k then Stop, else go to Step *Iteration*.

Due to the consideration of indicator vectors δ^m, the treatment of each node m at step k involves up to $\mathrm{O}(\overline{A}^+_m \overline{N})$ elementary operations, yielding a complexity in $\mathrm{O}(\overline{N}^2)$ by iteration hence of $\mathrm{O}(\overline{N}^3)$ overall.

Figure 3 provides an instance of circuit that arises if no caution is exerted: after a cost of 11.4 is obtained at n (cf. Fig. 2), using link e would yield a cost of 11.5 at node m, thus inducing a track cost of 11.7 for link d and reducing the cost at node n to $.2\times 9 + .8\times 11.7 = 11.16\ldots$ if loops are admitted then the limit cost at n satisfies that $c_n = .2\times 9 + .8\times(c_n + .3)$, yielding $c^\infty_n = 9.12$.

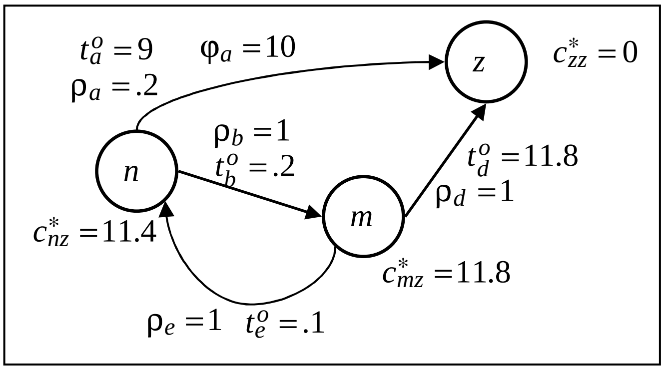

Fig. 3. Potential occurrence of loop.

As for a Dijkstra-like algorithm, Spiess and Florian [SpF89] have designed a shortest hyperpath algorithm that deals with links rather than nodes: starting from the destination, each link $a \approx (n,m)$ is dealt with when its current cost $c^o_{az} \equiv t^o_a + c_{mz}$ is minimal among a waiting list. Then, if $c^o_{az} < c_{nz}$ the provisional cost at n, the link is included in the local strategy at its tail node, yielding improved cost $c'_{nz} < c_{nz}$. Furthermore, the links b heading to n will be included in the waiting list with cost $t^o_b + c'_{nz}$.

This scheme can be adapted to the availability conditions by using auxiliary variables C, R, B, F, T and r at each node to perform the optimization of the local travel strategy as in Section 3.3. In the network algorithm, the node problems are addressed in parallel rather than sequentially. The loop avoidance must also be dealt with.

Yet this adaptation will have two additional differences from the Spiess-Florian algorithm. First, the minimum cost among the waiting list may not decrease at some steps, rather than keeping decreasing. This may happen typically when link a under examination provides a better recourse option to its tail node, thereby reducing its provisional cost below the track cost c_{az}^{o}: then if n provides access to a link b of very low cost, the track cost $t_b^o + c'_{nz}$ may be lower than c_{az}^{o}. Second, although the evaluation of a node cost in a basic hyperpath may involve several track costs, each of them is consolidated once only, whereas under availability track cost revision may change the rank of some tracks in an ordered sequence strategy under construction, thereby requiring to re-evaluate cumulated probabilities and provisional cost. Overall, the algorithm à la Dijkstra-Spiess-Florian for shortest hyperpaths under availability is merely a depth-first, cost-oriented exploration of the network. Its computational performance may not be better than that of the algorithm à la Ford-Bellman. Lastly, due to the lack of cost monotonicity along a hyperpath under availability, the standard property that a node cannot be re-examined if its average cost is less than that of any waiting track c_{az}^{o} is lost. It is replaced by the following adaptation: a node m with *recourse* cost T_m less than that of any waiting track has been optimized – because no track can improve on T_m hence on C_m since $T_m \geq C_m$.

5. Traffic equilibrium

So far the system conditions have been taken as exogenous in order to focus on the microeconomic routing behavior of a single network user. As there is a large set of network users, by aggregation their individual hyperpaths induce a macroscopic pattern of usage at both levels of origin-destination pairs and of network elements. By network element, the demanded flow induces the local quality of service; in turn, the local characteristics of supply induce the quality of service by option of route choice i.e. by hyperpath from origin to destination along the network.

These dependencies are depicted in figure 4, which shows that they constitute a circuit of interactions that is characteristic of some form of equilibrium between supply and demand.

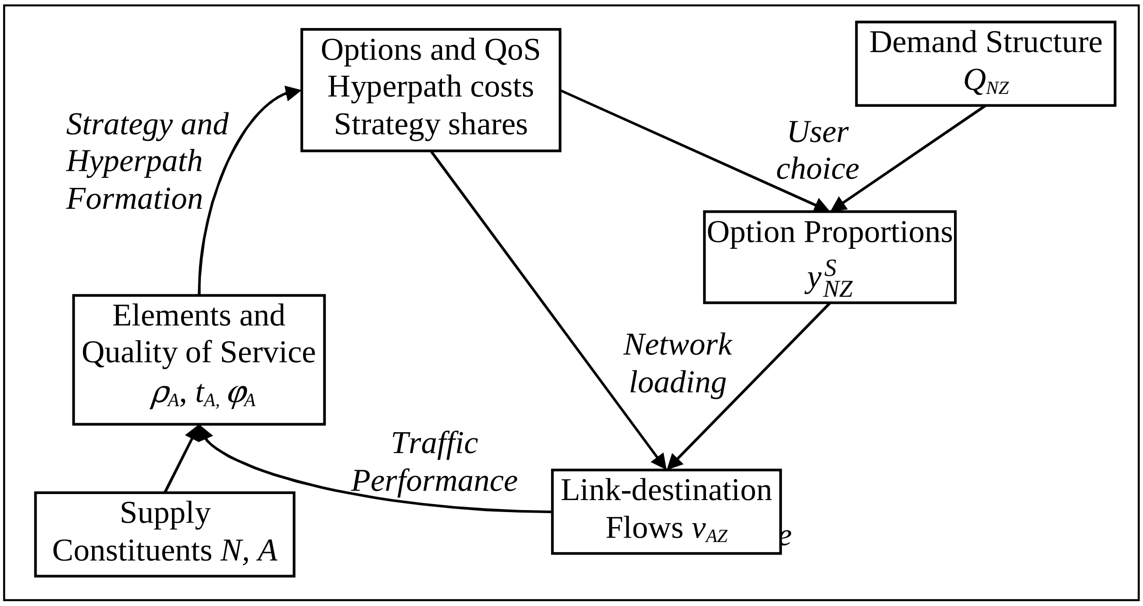

Fig. 4. Systemic analysis of demand and supply interactions.

Notation for traffic equilibrium

Z set of destination nodes $z \in N$

$q_{NZ} = [q_{nz} : n \in N, z \in Z]$ vector of origin-destination flow rates

$S(n)$ for $n \in N$ set of travel strategies s from node n

$y_{NZ}^{S} = [y_{nz}^{s} : z \in Z, n \in N, s \in S(n)]$ with y_{nz}^{s} proportion of flow from n to z assigned to s

$v_{AZ} = [v_{az} : a \in A, z \in Z]$ vector of link-destination flow rates

5.1 Traffic performance function

Let us take vectors of local traffic conditions, namely $t_A^o = [t_a^o : a \in A]$, ρ_A and φ_A as functions of the vector of flows by link and destination, $v_{AZ} = [v_{az} : a \in A, z \in Z]$. Each elementary function is assumed continuous with respect to v_{AZ}. The range of a t_a^o function is $\Re^+$ whereas that of ρ_a is $]0,1]$ and that of φ_a is $]0,\infty[$. The precise definition of each function must be consistent with the transportation mode of link a, cf. section 2. A private arc has constant $\rho_a = 1$ and its frequency can be taken as infinite but does not play any role. The functions ρ_a and φ_a are basically purported to characterize a link of public transportation, especially so for transit boarding links. Auxiliary functions can be useful to represent station dwell time, such as $d_a(v_{AZ})$ by vehicle run at the stop serviced by the transit line of a, and the residual capacity, $k_a(v_{AZ})$. Dwell time d_a may be related to the number of boarding and alighting passengers by vehicle run, which involves v_a, φ_a and also the flow on the alighting link at the same stop. It will also be limited by the operator to a maximum value so as to maintain the pace of operations. These considerations lead to the following function of local availability, in which $\rho_a^o < 1$: $\rho_a \equiv \max\{\rho_a^o, d_a \varphi_a / H\}$.

However the deepest consequence of modeling the PIA of a public mode is to capture saturated regimes simply, by letting ρ_a tend to 1 when the trip flow rate v_a approaches the capacity rate κ_a of service a. Such physical behavior of the PIA involves the following requirements:

i) at small v_a / κ_a i.e. uncongested regime, the PIA should conform to the basic definition, $\rho_a^{\text{free}} \equiv \max\{\rho_a^o, d_a \varphi_a / H\}$.

ii) When $v_a / \kappa_a \to 1$, PIA $\rho_a \to 1$ continuously at a quick pace. Moreover the uncongested regime should extend over a large range.

iii) The average cost of any strategy s including option a with positive share should increase smoothly with v_a. This requires to adapt time t_a^o of the boarding link along with ρ_a, since were t_a^o constant then the average cost would decrease with ρ_a.

To meet the second requirement, the following function of $\xi = v_a / \kappa_a$ can be used as default formula, using a parameter $\xi_a^\kappa \in]0,1[$: $\rho_a^{\text{sat}} = \min\{1, \rho_a^o + (1-\rho_a^o).(\xi - \xi_a^\kappa)^+ / (1-\xi_a^\kappa)\}$. Then a more general function meeting the two first requirements would be $\rho_a = \max\{\rho_a^{\text{free}}, \rho_a^{\text{sat}}\}$.

Alternative models of deeper physical significance can be specified by considering the residual capacity, which depends on the service frequency, the vehicle nominal capacity and its current occupation by on-board passengers. Taking vehicle residual capacity k_a as a continuous, strictly positive function of v_{AZ} and assuming a number φ_a of runs by homogeneous vehicles during H, the associated service has capacity rate of $\kappa_a = k_a \varphi_a / H$. On a boarding link it interacts with the flow rate of boarding candidates, $v_a = \sum_{z \in Z} v_{az}$ to yield waiting conditions. Assuming a local traffic bottleneck, incoming trip rate of v_a faced to capacity rate of κ_a during period H yields an average wait time of $\tilde{w}_a = \frac{1}{2} H(v_a / \kappa_a - 1)^+$ by passenger. Under saturation the persistent stock of waiting passengers is $N_a = \tilde{w}_a v_a$ (by Little's law) and $\tilde{\rho}_a = 1$ if this is higher than vehicle capacity k_a, so an ad hoc formula is:

$$\rho_a^{\text{sat}} \equiv \min\left\{1, \frac{Hv_a}{2k_a}\left(\frac{Hv_a}{\varphi_a k_a} - 1\right)^+\right\}.$$

This more elaborate model of PIA ρ_a^{sat} together with the bottleneck wait time $\tilde{w}_a$ cannot meet the third requirement, because $\tilde{w}_a + (1-\rho_a)H/\varphi_a$ will decrease not increase as v_a approaches κ_a. More generally, if a belongs to strategy s at position $i+1$ then its contribution to the average strategy cost amounts to $\overline{R}_i^s(\rho_a t_a^o + (1-\rho_a)\bar{t})$, in which $\bar{t}$ denotes a combined cost that may involve t_a^o in the recourse part. Given $\bar{t}$, $\rho_a(t_a^o - \bar{t})$ increases with ρ_a if (cf. Appendix)

$$t_a^o = t_a^o\big|_{\rho_{a0}} + \frac{H}{\varphi_a}\left(\frac{\rho_a}{\rho_{a0}} - 1\right) \text{ for } \rho_a \geq \rho_{a0}.$$

So it is convenient to set

$$\hat{w}_a \equiv t_a^o\big|_{\rho_a} - t_a^o\big|_{\rho_{a0}} = \frac{H}{\varphi_a}\left(\frac{\rho_a}{\rho_{a0}} - 1\right)^+.$$

Taking $\hat{w}_a$ as minimum time t_a^o of the boarding link amounts to add it to the minimum time of the service line. In the uncongested regime there is no effect since $\rho_a \leq \rho_{a0}$. Near capacity the effect is to add one or several service headways to the user's service time, meaning that the user has no chance to board the first incoming vehicle. This is consistent because if $v_a = \kappa_a$ then a queue of waiting passengers is persistent and obviates any immediate boarding from a newly coming user.

The consequences on route choice are illustrated in section 6.2.

5.2 Demand behavior and the propagation of flow

Every network user is assumed to behave rationally by using local travel strategies of minimal cost. Between node n and destination z, the expected cost of a strategy $s \in S(n)$ is a function

$$g_{nz}^s = \mathrm{C}_{nz}(s, [t_a^o + g_{mz}^*, \rho_a, \varphi_a : a \approx (n,m) \in A_n^+]), \tag{9}$$

In which g_{mz}^* is the minimum downstream cost from m to z. The formation of costs to destination z from all nodes n is determined by conditions (7) that make a set of generalized Bellman equations, for which a solution has been shown to exist and to be unique in costs (Lemma 1).

At node n, let y_{nz}^{s} denote the proportion of unit assigned to strategy s. User behavior amounts to the following conditions (Wardrop principle for travel strategies):

$$g_{nz}^{*} = \min_{s \in S(n)} g_{nz}^{s} \tag{10a}$$

$$y_{nz}^{s} \geq 0\,,\ \sum_{s \in S(n)} y_{nz}^{s} = 1 \tag{10b}$$

$$y_{nz}^{s}.(g_{nz}^{s} - g_{nz}^{*}) = 0\,. \tag{10c}$$

Concerning the propagation of flows, the flow x_{nz} coming in n and destined to z stems from local inflow q_{nz} and the link flows destined to z that come from upstream, i.e.

$$x_{nz} = q_{nz} + \sum_{b \in A_n^-} v_{bz}\,. \tag{11}$$

Furthermore, the assignment of network flows to local strategies imposes that

$$v_{az} = x_{nz} \sum_{s \in S(n)} y_{nz}^{s} \pi_a^{s}\,, \tag{12}$$

in which π_a^s is the market share that depends solely on the link characteristics. This last dependency can be stated formally on the basis of a function denoted P_A^S as follows:

$$(\rho_A, t_A, \varphi_A) \mapsto \pi_A^S = [\pi_a^s : \forall n \in N, \forall a \in A_n^+, \forall s \in S(n)]\,. \tag{13}$$

5.3 Statement of traffic equilibrium

The system state is a pair of vectors, $[v_{AZ}, y_{NZ}^S]$ made of the vector of link-destination flow and the vector of strategy assignments. Denoting $v_A \equiv \sum_{z \in Z} v_{Az}$, let us then define a traffic equilibrium as a vector pair $[v_{AZ}, y_{NZ}^S]$ that satisfies the set of conditions as follows:

- Traffic performance: $t_A^o = \mathrm{T}_A^o(v_A)$, $\rho_A = \mathrm{R}_A(v_A)$ and $\varphi_A = \mathrm{F}_A(v_A)$, in which the vector functions T_A^o, R_A and F_A are continuous and strictly positive.
- Hyperpath costs: (9) and (10a).
- User optimality: (10b) and (10c).
- Flow conservation: (11), (12) and (13).

Such a set of conditions is fairly standard for capacitated static equilibrium of traffic on a transit network (e.g. [CoC01], [CCF06]). The availability framework induces specific traffic performance functions and local strategies.

5.4 Characterization and existence of traffic equilibrium

Lemma 2. *For each (t_A, ρ_A, φ_A) there is a unique solution g_{NZ}^{*} of the system of shortest hyperpaths conditions (Bellman conditions), which depends continuously on (t_A, ρ_A, φ_A).*

Proof. The existence and uniqueness have been established in Lemma 1 and Proposition 4, yielding g_{NZ}^{*} as a function $G_{NZ}(t_A, \rho_A, \varphi_A)$. Continuity is established by straightforward adaptation to the availability framework of Proposition 1 in [CoC01] as follows. Define by (9) and (10a) applied *at any single stage*, a mapping $G_{NZ}^{*} = [G_{nZ}^{*} : n \in N]$ with respect to $[g_{NZ}^{*}, t_A, \rho_A, \varphi_A]$: the argument g_{NZ}^{*} corresponds to the g_{mz}^{*} in the right hand side of (9).

This elementary, local mapping is continuous on its domain since it involves only operations of strategy cost evaluation and of minimization, both of which are continuous. Let now (t^n, ρ^n, φ^n) be a sequence converging to (t_A, ρ_A, φ_A) and let $g^{*n}_{NZ} = G_{NZ}(t^n, \rho^n, \varphi^n)$ be the associated solutions of the generalized Bellman equations. Any accumulation point $\tilde{g}$ in this sequence satisfies that $g^{*n(j)}_{NZ} = G_{NZ}(t^{n(j)}, \rho^{n(j)}, \varphi^{n(j)}) \to \tilde{g}$ for a subsequence $n(j)$ so that, by stating local conditions and by letting $j \to \infty$, it comes out that $\tilde{g} = G^*_{NZ}(\tilde{g}, t, \rho, \varphi)$: then, by the uniqueness of the GBE solution, it must hold that $\tilde{g} = G(t, \rho, \varphi)$. Still due to uniqueness, the latter condition implies that there is a unique accumulation point $\tilde{g}$ in the sequence (t^n, ρ^n, φ^n), which makes it its limit as the sequence belongs to a compact set. Therefore the function G_{NZ} is continuous with respect to (t_A, ρ_A, φ_A).

Let $E \equiv \left\{ y^S_{NZ} = [y^s_{nz} : n \in N, z \in Z, s \in S(n)] \text{ such that } y^s_{nz} \geq 0 : \sum_{s \in S(n)} y^s_{nz} = 1 \right\} \subset [0,1]^{S(N) \times Z}$ be the set of strategy assignments by node and destination. It is a convex polytope of finite dimension, hence a compact set.

The flow vector by destination and link, v_{AZ}, is related to the vector of strategy assignments, y^S_{NZ}, on the basis of (11) and (12). These also involve the O-D matrix of trip flows, q_{NZ}, and the strategy modal shares, π^S_A. Given q_{NZ}, let us check that (11) and (12) constitute a functional relationship linking v_{AZ} to y^S_{NZ} and π^S_A, denoted hereafter V_{AZ}.

By destination z, from each node $n \neq z$, let us denote $h_{az} \equiv \sum_{s \in S(n)} y^s_{nz} \pi^s_a$ and define $h^z_{mn} \equiv \sum_{a \in A^+_m \cap A^-_n} h_{az}$ the coefficient of flow transfer from node m to node n. As the (y^s_{nz}) constitute a distribution of probability among the strategies $s \in S(n)$ and so do the (π^s_a) among the links $a \in s$, so do the (h_{az}) among the links emanating from node m and the (h^z_{mn}) among the node-to-node transitions tailed at m. By adding (12) over the links heading from m to n, the $[x_{nz} : n \in N]$ must satisfy that $\sum_{a \in A^+_m \cap A^-_n} v_{az} = x_{mz} h^z_{mn}$; substituting in (11) yields that $x_{nz} = q_{nz} + \sum_{m \in N} x_{mz} h^z_{mn}$. In vector form, $x_{NZ} = q_{NZ} + x_{NZ} \mathbf{h}^z$ or equivalently, letting $\mathbf{u}_N$ denote the identity matrix on $N \times N$,

$$x_{NZ}(\mathbf{u}_N - \mathbf{h}^z) = q_{NZ}. \tag{14}$$

As matrix $\mathbf{h}^z$ contains only positive terms, so does $\mathbf{m}^z_I \equiv \sum_{i=0}^{I} (\mathbf{h}^z)^i$ for every $I \geq 0$. This matrix satisfies that $(\mathbf{u}_N - \mathbf{h}^z).\mathbf{m}^z_I = \mathbf{u}_N - (\mathbf{h}^z)^{I+1}$. If the strategy proportions are associated to a hyperpath, then for $i \geq \overline{N}$ it holds that $(\mathbf{h}^z)^i = 0$ so that $\mathbf{m}^z_{\overline{N}}$ stands as the inverse matrix of $\mathbf{u}_N - \mathbf{h}^z$, making x_{NZ} a continuous function of q_{NZ} and (y^S_{NZ}, π^S_A). Then, as v_{AZ} stems from x_{NZ} owing to (12), mapping V_{AZ} is a function, and a continuous one.

So, by construct, function V_{AZ} is continuous. So is function P^S_A defined in (13) which relates π^S_A to (ρ_A, t_A, φ_A), and function $\mathrm{U}_A : v_{AZ} \mapsto v_A \equiv \sum_{z \in Z} v_{Az}$. By sequential composition of continuous mappings, mapping $\tilde{\mathrm{C}}^S_{NZ} \equiv \mathrm{C}^S_{NZ} \circ (R_A, T_A, F_A)$ is continuous with respect to (y^S_{NZ}, v_A), as well as $\tilde{\mathrm{G}}^S_{NZ} \equiv \tilde{\mathrm{C}}^S_{NZ} \circ \mathrm{U}_A \circ \mathrm{V}_{AZ}$ with respect to (y^S_{NZ}, v_{AZ}).

Proposition 5. *A system state $(y^{*S}_{NZ}, v^{*}_{AZ})$ made of a strategy assignment y^{*S}_{NZ} and a network flow state v^{*}_{AZ} is a traffic equilibrium if and only if it satisfies the following conditions: letting $\pi^{*S}_{A} \equiv \mathrm{P}^{S}_{A} \circ (R_A, T_A, F_A) \circ \mathrm{U}_A(v^{*}_{AZ})$,*

$$v^{*}_{AZ} = \mathrm{V}_{AZ}(y^{*S}_{NZ}, \pi^{*S}_{A}), \tag{15a}$$

$$\forall y^{S}_{NZ} \in E,\ (y^{S}_{NZ} - y^{*S}_{NZ}).\tilde{\mathrm{G}}^{S}_{NZ}(y^{*S}_{NZ}, v^{*}_{AZ}) \geq 0. \tag{15b}$$

Proof. The definition of π^{*S}_{A} and condition (15a) ensure that the flow vector satisfies the traffic performance conditions as well as those of flow conservation in the definition of traffic equilibrium. Moreover, the Variational Inequality Problem (VIP) condition (15b) implies that

$$\forall z \in Z, \forall n \in N, \forall s \in S(n),\ \ y^{*s}_{nz} > 0\ \Rightarrow \tilde{\mathrm{G}}^{s}_{nz}(y^{*S}_{NZ}, v^{*}_{AZ}) = \min_{\sigma \in S(n)} \tilde{\mathrm{G}}^{\sigma}_{nz}(y^{*S}_{NZ}, v^{*}_{AZ}), \tag{16}$$

which amounts to the Wardropian condition of user optimality of a local travel strategy.

Thus, by construct of the $\tilde{\mathrm{G}}^{S}_{NZ}$ mapping, all of the equilibrium conditions are satisfied jointly, which implies that the pair y^{*S}_{NZ}, v^{*}_{AZ} is a traffic equilibrium. Conversely, if the VIP condition (15b) is not satisfied then at least one of the equilibrium conditions (16) does not hold.

Proposition 6. *If the traffic performance functions are continuous, then there exists a solution to the problem of traffic equilibrium (15).*

Proof. Under given proportions π^{S}_{A}, (15a) can be embedded into $\tilde{\mathrm{G}}^{S}_{NZ}$ to make it a function of y^{S}_{NZ} only, and a continuous one indeed if the traffic performance functions are continuous. Then, as a VIP with continuous cost function in a non-empty, compact set, (15b) has a solution which is a function of π^{S}_{A}. Uniqueness and regularity can be ensured by requiring further solution minimality according to a given norm. Then v^{*}_{AZ} also is a regular function of π^{S}_{A}, denoted $\tilde{\mathrm{V}}_{AZ}(\pi^{S}_{A})$, so that the problem of traffic equilibrium (15) amounts to solve a fixed point problem with respect to π^{S}_{A} and mapping $\mathrm{P}^{S}_{A} \circ (R_A, T_A, F_A) \circ \mathrm{U}_A \circ \tilde{\mathrm{V}}_{AZ}$ which is continuous. As the domain is a nonempty compact set, there must be a solution to that fixed point problem, which induces a traffic equilibrium.

5.5 A heuristic algorithm

It would be computationally awkward to deal with strategy proportions y^{s}_{nz} explicitly, since there is a combinatorial number of them at each node for each destination. It is much simpler to focus on the associated network flow state by link and destination, v_{AZ}, and to ensure that the equilibrium conditions (15) are satisfied.

A Method of Successive Averages (MSA) is an adequate though heuristic algorithm to solve for equilibrium. Let us recall shortly that it involves a sequence of decreasing positive numbers $\xi_k \in]0,1[$ for $k > 0$ and $\xi_0 = 1$ together with a current state v_{AZ} and an auxiliary state $\tilde{v}_{AZ}$ as follows:

Initialization. Let $k := 0$ and $v_{AZ} := 0$.

Traffic performance. Let $v_A := \sum_{z \in Z} v_{Az}$, $t^{o}_{A} := \mathrm{T}^{o}_{A}(v_A)$, $\rho_A := \mathrm{R}_A(v_A)$ and $\varphi_A := \mathrm{F}_A(v_A)$.

Auxiliary state. For every destination z, find the optimal hyperpath from each node to z on the basis of the local conditions, and assign the origin-destination flows to optimal travel strategies, yielding $\tilde{v}_{AZ}$.

Convex combination. Update $v_{AZ} := (1-\xi_k)v_{AZ} + \xi_k \tilde{v}_{AZ}$.

Termination test. If v_{AZ} and k satisfy a convergence criterion then Stop, else update $k := k+1$ and go to Step *Traffic Performance*.

The convex combination on v_{AZ} implies an underlying convex combination on the underlying strategy proportions y_{NZ}^{S}, though with coefficients distinct by node and destination since they are weighted by the node flow $x_{nz} = q_{nz} + \sum_{b \in A_n^-} v_{bz}$:

$$y_{nz}^{s(k+1)} := [(1-\xi_k)x_{nz}^{(k)} y_{nz}^{s(k)} + \xi_k \tilde{x}_{nz}^{(k)} \tilde{y}_{nz}^{s(k)}] / [(1-\xi_k)x_{nz}^{(k)} + \xi_k \tilde{x}_{nz}^{(k)}].$$

5.6 A rigorous convergence criterion

As for a convergence criterion, it must involve the two equilibrium conditions (15a,b). The part of the criterion dealing with user optimality (15b) should be based on the strategy proportions. This leads us to the problem of recovering strategy proportions from link flows by destination. In the congestible strategy model of [CCF06], the approach is to decompose link flows (at a given node and to a given destination) with respect to strategies in a progressive way: at a given step, the strategy that involves all links with strictly positive residual flow is factorized by the largest possible coefficient and deduced from the residual link flows, yielding updated residuals with strictly smaller link supporting set. Then function $\sum_{s \in S(n)} y_{nz}^{s}(g_{nz}^{s} - g_{nz}^{*})$ is evaluated as an upper bound, which reduces to zero when only optimal strategies are used. Overall, function $\mathrm{J}_y \equiv \sum_{n,z,s} y_{nz}^{s}(g_{nz}^{s} - g_{nz}^{*})$ is a non negative objective function, of which a null value reveals that condition (15b) holds. Furthermore, [CCF06] have provided an evaluation formula for that function that only depends on the flows by link and destination.

Unfortunately, such an explicit formula is unlikely to exist in the availability framework: therefore we shall restrict ourselves to a less strict upper bound as follows.

Therefore, we introduce hereafter an algorithm to obtain an upper bound: given destination z and node n, let $v = [v_{az} : a \in A_n^+]$ denote the sub-vector of flows by link going out of n and by destination z. While $v \neq 0$, denote by s'_v the strategy subset of s_n using only links such that $v_{az} > 0$ and find the optimal strategy s' within s'_v: denote by π' the routing proportions in s' and let $r := \min\{v_{az} / \pi'^{s'}_a : a \in s'\}$. The residual vector $v - r\pi' \geq 0$ with at least one more null component than v. Replace v by $v - r\pi'$ and repeat until the residual flow is zero. By construction, $v = \sum_\ell r_\ell \pi'^{s(\ell)}$ and the following indicator $\mathrm{DG}_{nz} := \sum_\ell r_\ell (g_{nz}^{s(\ell)} - g_{nz}^{*})$ is an estimator of convergence to optimality for that pair of destination and current node. Furthermore, if all of the selected strategies are optimal then the gap indicator reduces to zero.

The computational complexity amounts to $\mathrm{O}(\overline{A_n^+}^2 \ln \overline{A_n^+})$ since at most $\overline{A_n^+}$ steps ℓ are required, each one with complexity $\mathrm{O}(\overline{A_n^+} \ln \overline{A_n^+})$. So the evaluation is more expensive than for the upper bound in [CCF06], but still quite affordable as it will only add a factor $\max_{n \in N} \overline{A_n^+}$ on the overall complexity of the equilibration algorithm.

Along with the VIP part (15b), the fixed point part (15a) about v_{AZ} should be checked by the convergence criterion, too. In fact the local decomposition of v_{AZ} to recover strategy proportions y_{NZ}^S ensures that v_{AZ} is consistent with those y_{NZ}^S and the current π_A^S. If the latter are derived from v_{AZ} on the basis of $\pi_A^S = \mathrm{P}_A^S \circ (R_A, T_A, F_A) \circ \mathrm{U}_A(v_{AZ})$, and if $\mathrm{J}_y \to 0$, then conditions (15a) are satisfied implicitly, so that J_y is a rigorous convergence criterion for the whole problem of traffic equilibrium.

6. Numerical instance

This section addresses a small instance in order to demonstrate two features of the availability model: first, the user choice between alternative options on the basis of travel strategies; second, the effect of service saturation on the option shares within a strategy, then in turn on traffic equilibrium.

Fig. 5 depicts the toy network: one O-D pair from origin node o to destination node z, serviced by three options among which two are transit services (links a, b) and the other one a Walk mode (link w). Nine travel strategies are available from o, namely: the single options $\{a\}$, $\{b\}$, $\{w\}$, the pairs $\{a,b\}$, $\{b,a\}$, $\{a,w\}$, $\{b,w\}$ and the triples $\{a,b,w\}$ and $\{b,a,w\}$.

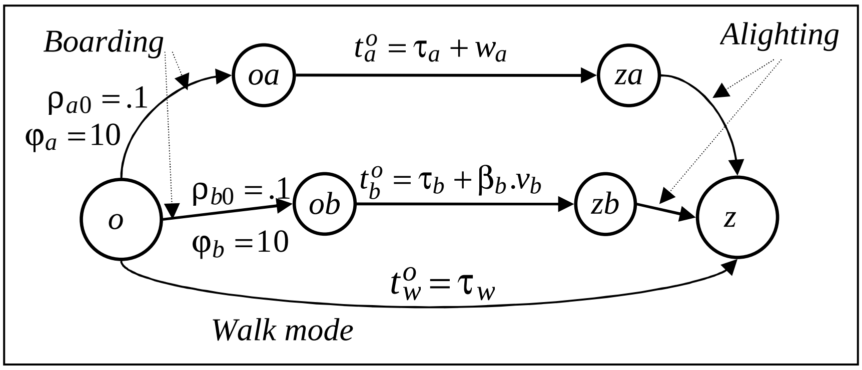

Fig. 5. Toy network.

6.1 Uncapacitated assignment

Let us take all the traffic conditions as in Fig. 5, except for a varying travel time along route b that depends on the option flow assigned to it, v_b plus an exogenous preload ξ as follows: $t_b^o = \tau_b + \beta_b(v_b + \xi)$.

Taking an OD flow rate of $q = 100$ trips per hour, $\tau_a = \tau_b = 10$', $\beta_b = 0.1$ min.h/trip and $\tau_w = 20$', the variations of parameter ξ ranging from 0 to 1,000 trips per hour yields flow assignment and strategy times depicted in figures 6 and 7, respectively. In this experiment only the transit options will be used, since the walk mode is not competitive. On the interval $\xi \in [0,555]$ the ordered sequence $\{a,b\}$ is optimal; the second best strategy $\{b,a\}$ delivers very close performance but gets no flow at all. When $\xi \in [555,600]$ strategies $\{a,b\}$ and $\{a\}$ are efficient jointly. When $\xi > 600$ the mixed strategy is outperformed by the single option $\{a\}$ to carry flow from o to z.

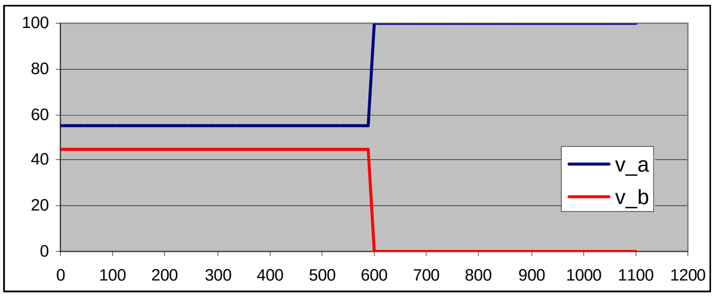

Fig. 6. Assignment of flow to options depending on parameter ξ.

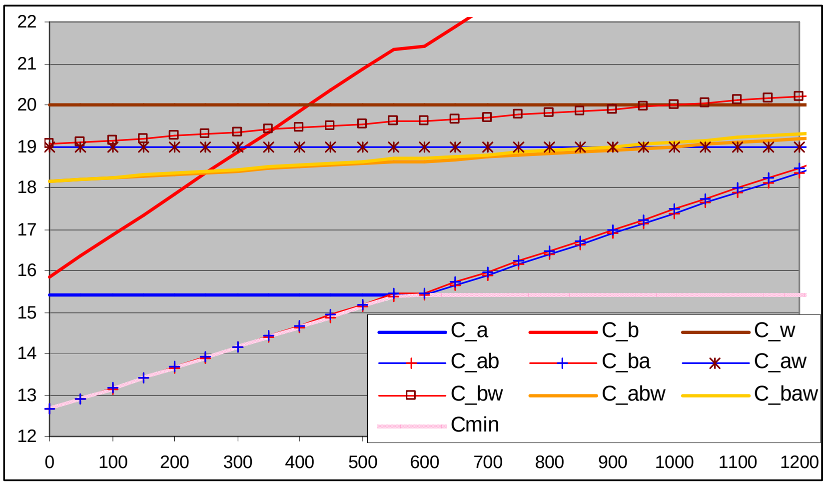

Fig. 7. Strategy cost depending on parameter ξ.

6.2 Capacitated assignment

Assume now that there is no preload nor capacity limit on line b, whereas transit route a is serviced by vehicles of capacity $k_a = 50$ passengers. Then the capacity rate of line a amounts to $\kappa_a \equiv k_a \varphi_a / H = 500$ trips per hour. Here the capacity model involves PIA ρ_a with respect to $\xi = v_a / \kappa_a$ and boarding wait $\hat{w}_a$, both of which are taken from section 5.1 as follows:

$$\rho_a = \min\{1, \rho_a^o + (1-\rho_a^o).\frac{(\xi - \xi_a^\kappa)^+}{1-\xi_a^\kappa}\} \text{ wherein } \rho_a^o = d_a \varphi_a / H \text{ and } \xi_a^\kappa = 90\%,$$

$$\hat{w}_a = \frac{H}{\varphi_a}\left(\frac{\max\{\rho_a, \rho_a^o\}}{\rho_a^o} - 1\right).$$

The variation of OD flow rate q_{oz} ranging from 0 to 2,000 trips/hour yields flow assignment and strategy costs depicted in figures 8 and 9, respectively.

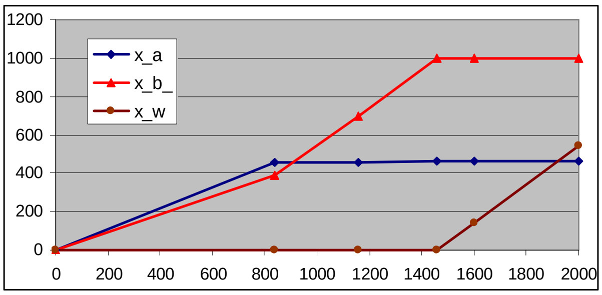

Fig. 8. Assignment of flow to options depending on parameter q_{oz}.

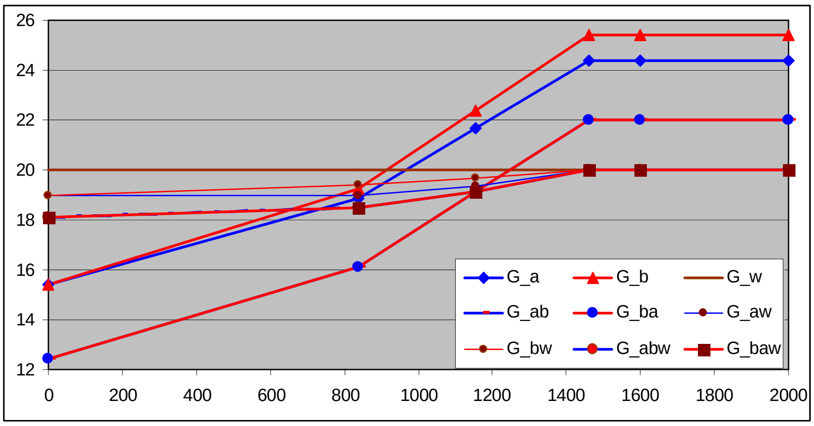

Fig. 9. Strategy cost depending on parameter q_{oz}.

When the OD flow is low, both transit links are attractive. Among them, line a is more competitive since its minimum time is given while that of line b increases with its link volume. Consequently, the ordered sequence $\{a,b\}$ is optimal; whereas strategy $\{b,a\}$ is second best only despite its average cost is very close. When v_a approaches κ_a, then t_a^o including the wait time prior to boarding increases quickly; it attains t_b^o at value q_{ab}^* of the OD flow, at which point $\{a,b\}$ and $\{b,a\}$ have equal average cost. This remarkable value is characterized in the appendix and evaluated at $q_{ab}^* \approx 839$ p/h. Then, for q beyond q_{ab}^*, both strategies are equivalent, with same average cost and also same recourse cost. They are optimal until their recourse cost attains the cost by walking, t_w^o. This takes place at remarkable value q_{abw}^* such that $t_w^o = \frac{H}{\varphi_a+\varphi_b} + t_b^o$, also knowing that $t_b^o = t_a^o$. There, x_a^* and x_b^* stem from their respective travel time function, yielding that $q_{abw}^* = x_a^* + x_b^*$ = 1,156 p/h. From this value the optimal strategies are $\{a,b,w\}$ and $\{b,a,w\}$, until $t_b^o = t_a^o = t_w^o$ at the last remarkable value q_w^* at which point the discrete strategy $\{w\}$ becomes attractive. Recovering x_a^* and x_b^* from their respective travel time function, it comes out that $q_w^* = x_a^* + x_b^*$ = 1,459 p/h. Above this value, the five strategies $\{w\}$, $\{a,w\}$, $\{b,w\}$, $\{a,b,w\}$ and $\{b,a,w\}$ are efficient, with average cost that does not vary any longer.

7. Conclusion

To deal with the immediate availability of travel options explicitly, a comprehensive framework has been provided, including the basic model of a travel option, the local model of route choice and travel strategies, the network model with sequential composition of paths and traffic equilibrium. The availability characteristics extend the notion of travel strategy. It diminishes the role of the Markovian assumption about public services. Furthermore, some features of capacity constraints are easy to model in the probability of immediate availability.

The availability issue is relevant not only to mass transit but also to individual public modes such as vehicle-sharing services (VSS): such modes involve a dual notion of availability, namely vehicle availability vs. locker availability, which necessitates to elaborate the availability model (cf. [Leu13]).

These refinements in the representation of travel modes come at the price of an additional computational burden. Concerning the local problem of optimal strategy, the computational complexity both in time and space amounts to about twice that of the traditional common line problem. The relative cost of a hyperpath search may be somewhat higher, mostly due to the need to detect and avoid circuits. In principle, the search for an equilibrium state should entail no additional cost compared to traditional equilibrium transit assignment models: however the evaluation of the convergence criterion is computationally more costly, with relative factor equal to the average node out-degree (the number of links going out of a node).

8. References

[CCF06] Cepeda M., R. Cominetti and M. Florian, A frequency-based assignment model for congested transit networks with strict capacity constraints: characterization and computation of equilibria, *Transportation Research Part B*, Vol. 40, 2006, pp. 437–459.

[ChR75] Chriqui C. and P. Robillard, Common bus lines, *Transportation Science*, Vol. 9, 1975, pp. 115–121.

[CoC01] Cominetti R. and J. Correa, Common-lines and passenger assignment in congested transit networks, *Transportation Science*, Vol. 35, 2001, pp. 250–267.

[HMN04] Hamdouch Y., P. Marcotte and S. Nguyen, Capacitated transit assignment with loading priorities, *Mathematical Programming B*, Vol. 101, 2004, pp. 205–230.

[KBS03] Kurauchi F., M. G. H. Bell and J.D. Schmöcker, Capacity constrained transit assignment with common lines, *Journal of Mathematical Modelling and Algorithms*, Vol. 2, 2003, pp. 309-327.

[Kle75] Kleinrock L. *Queueing systems*. Volume I: Theory. Wiley, New York, 1975.

[LCP12] Leurent F., E. Chandakas and A. Poulhès. A Passenger Traffic Assignment Model with Capacity Constraints for Transit Networks. In *Elsevier Procedia - Social and Behavioral Sciences*, Volume 54, 4 October 2012, Pages 772–784

[LeC12] Leurent F. and E. Chandakas. The Transit Bottleneck Model. In *Elsevier Procedia - Social and Behavioral Sciences*, Volume 54, 4 October 2012, Pages 822–833.

[Leu06] Leurent F. *Structures de réseaux et modèles de cheminement*. Ed. Lavoisier, Paris, 2006.

[Leu13] Leurent F. Travel Demand for a One-Way Vehicle Sharing System: a model of traffic assignment to a multimodal network with supply-demand equilibrium. Abstract submitted to the MT-ITS Conference, March 2013.

[NgP88] Nguyen S. and S. Pallotino, Equilibrium traffic assignment in large scale transit networks, *European Journal of Operational Research*, Vol. 37, 1988, pp. 176–186.

[SpF89] Spiess H. and M. Florian, Optimal strategies : a new assignment model for transit networks, *Transportation Research Part B*, Vol. 23, 1989, pp. 83–102.

9. Appendix on capacity model

9.1 PIA and unavoidable wait time at boarding

For a boarding link a, let us study the relationship between the specific wait time t_a^o for that link that comes in addition to the wait time for the next vehicle arrival. Given an alternative time, $\bar{t}$, requirement (iii) in Section 5.1 is that $\rho_a(t_a^o - \bar{t})$ increases with ρ_a. Assuming differentiability, this amounts to $\frac{\mathrm{d}}{\mathrm{d}\rho_a}\rho_a(t_a^o - \bar{t}) \geq 0$ i.e. $t_a^o - \bar{t} + \rho_a \frac{\mathrm{d}t_a^o}{\mathrm{d}\rho_a} \geq 0$ and equivalently to

$\rho_a \frac{\mathrm{d}t_a^o}{\mathrm{d}\rho_a} \geq \bar{t} - t_a^o$

$\frac{\mathrm{d}t_a^o}{\bar{t}-t_a^o} \geq \frac{\mathrm{d}\rho_a}{\rho_a}$ as it can be expected that $\bar{t} \geq t_a^o$ for a competitive link

Therefore, $-\Delta\ln(\bar{t} - t_a^o) \geq \Delta\ln(\rho_a)$ between initial and final values ρ_a^o and $\rho_a \geq \rho_a^o$. This is equivalent to $\ln(\bar{t} - t_i) - \ln(\rho_a / \rho_a^o) \geq \ln(\bar{t} - t_f)$ and also to

$$t_f \geq \bar{t} - (\bar{t} - t_i)\frac{\rho_a^o}{\rho_a} \quad . \tag{A1}$$

Taking $t_a^o \equiv t_f$, $\bar{t} = t_f + H/\varphi_a$ and $t_i \equiv \tau_a$ in (A1) yields that $t_a^o \geq \tau_a + \frac{H}{\varphi_a}(\frac{\rho_a}{\rho_a^o} - 1)^+$.

Now, for conveniency, let us design a simple saturating PIA with respect to the ratio of flow to capacity, $\xi \equiv v_a / \kappa_a$. PIA can be kept constant at ρ_a^o until a limit ratio ξ_a^κ, for instance 90%. Then, between ξ_a^κ and 1, assume that the PIA varies linearly with $\xi - \xi_a^\kappa$, so that

$$\rho_a = \min\left\{1, \rho_a^o + \frac{1-\rho_a^o}{1-\xi_a^\kappa}.(\xi - \xi_a^\kappa)^+\right\}. \tag{A2}$$

In the numerical instance of Section 6.2, parameters values of $\rho_a^o = .1$ and $\xi_a^\kappa = .9$ are taken. The minimum travel time of line a is fixed at τ_a; the additional wait time makes t_a^o a piecewise affine function of v_a:

$$t_a^o = \tau_a + \frac{H(1-\rho_a^o)}{\varphi_a\rho_a^o(1-\xi_a^\kappa)}.(\frac{v_a}{\kappa_a} - \xi_a^\kappa)^+ = \tilde{\tau}_a + \beta_a(v_a - \xi_a^\kappa\kappa_a)^+$$

wherein $\beta_a \equiv \frac{H(1-\rho_a^o)}{\varphi_a\rho_a^o(1-\xi_a^\kappa)\kappa_a}$ and $\tilde{\tau}_a \equiv \tau_a - \beta_a\xi_a^\kappa\kappa_a$.

So if $\theta > \tau_a$ then condition $t_a^o = \theta$ yields the solution $v_a = \xi_a^\kappa\kappa_a + (\theta - \tilde{\tau}_a)/\beta_a$.

9.2 Derivation of remarkable values in section 6.2

When v_a approaches κ_a, then t_a^o including the wait time prior to boarding increases quickly; it attains t_b^o at value q_{ab}^* of the OD flow, at which point $\{a,b\}$ and $\{b,a\}$ have equal average cost. The condition that $t_a^o = t_b^o$ with common value denoted as θ implies that on line b, $v_b = (\theta - \tau_b)/\beta_b$. On combining the conditions derived for both links, we get a relationship between $q = v_a + v_b$ and θ as follows:

$$q = \xi_a^\kappa\kappa_a + \frac{\theta - \tilde{\tau}_a}{\beta_a} + \frac{\theta - \tau_b}{\beta_b} \text{ or equivalently, } \theta = \frac{\beta_a\beta_b(q - \xi_a^\kappa\kappa_a) - \beta_b\tilde{\tau}_a - \beta_a\tau_b}{\beta_a + \beta_b}.$$

This yields firstly q_{abw}^* by taking $\theta = \tau_w - H/(\varphi_a + \varphi_b)$ and secondly, q_w^* by taking $\theta = \tau_w$.

The derivation of q^*_{ab} is more intricate. This value comes out as the minimum OD flow q such that t^o_a attains t^o_b under strategy $s=\{a,b\}$. This can be stated as the minimum v_a such that $t^o_a(v_a)=t^o_b(\frac{\pi^s_b}{\pi^s_a}v_a)$. As $\pi^s_a=\rho_a+\overline{\rho}_a\overline{\rho}_b\varphi'_a$ wherein $\overline{\rho}_a\equiv 1-\rho_a$ and $\varphi'_a\equiv\varphi_a/(\varphi_a+\varphi_b)$, $\pi^s_b=1-\pi^s_a=\overline{\rho}_a(1-\overline{\rho}_b\varphi'_a)$ which we take as our unknown variable and denote by x.

As $\overline{\rho}_a=\overline{\rho}^o_a\dfrac{1-v_a/\kappa_a}{1-\xi^\kappa_a}$, it holds that $v_a=\kappa_a[1-\dfrac{\overline{\rho}_a}{\overline{\rho}^o_a(1-\xi^\kappa_a)}]$, hence that $v_a=\kappa_a[1-Mx]$ by letting $M\equiv\dfrac{1-\xi^\kappa_a}{\overline{\rho}^o_a(1-\overline{\rho}_b\varphi'_a)}$.

We can now use the condition $t^o_a(v_a)=t^o_b(\frac{\pi^s_b}{\pi^s_a}v_a)$ to derive a second order equation in x:

$\tilde{\tau}_a+\beta_a v_a=\tau_b+\beta_b\dfrac{x}{1-x}v_a$ hence $v_a[\beta_a(1-x)-\beta_b x]+(\tilde{\tau}_a-\tau_b)(1-x)=0$ or

$(1-Mx)[\beta_a-\beta_\Sigma x]+\dfrac{\tilde{\tau}_a-\tau_b}{\kappa_a}(1-x)=0$ or $M\beta_\Sigma x^2-(M\beta_a+\beta_\Sigma+\frac{\tilde{\tau}_a-\tau_b}{\kappa_a})x+\beta_a+\frac{\tilde{\tau}_a-\tau_b}{\kappa_a}=0$.

Numerically, $\beta_a=1.08$, $\tilde{\tau}_a=$-476 and $M=$.202, yielding solution $x=$.4596, from which q^*_{ab} = 839.41 is recovered.